Statistical Thermodynamics
for metaequilibrium or metastable states

A. Carati,[*] A. Maiocchi,[†] and L. Galgani[‡]

Department of Mathematics, Università degli Studi di Milano - Via Saldini 50, I-20133 Milano, Italy

We show how statistical thermodynamics can be formulated in situations in which thermodynamics applies, while equilibrium statistical mechanics does not. A typical case is, in the words of Landau and Lifshitz, that of partial (or incomplete) equilibrium. One has a system of interest in equilibrium with the environment, and measures one of its quantities, for example its specific heat, by raising the temperature of the environment. However, within the observation time the global system settles down to a state of apparent equilibrium, so that the measured value of the specific heat is different from the equilibrium one. In such cases formulæ for quantities such as the effective specific heat exist, which are provided by Fluctuation Dissipation theory. However, what is lacking is a proof that internal energy exists, i.e., that the fundamental differential form $\delta Q - \delta W$ (difference between the heat absorbed by the system and the work performed by it) is closed. Here we show how the coefficients of the fundamental form can be expressed in such a way that the closure property of the form becomes manifest, so that the first principle is proven. We then show that the second principle too follows, and indeed as a consequence of microscopic time–reversibility. The treatment is given in a classical Hamiltonian setting. One has a global time–independent Hamiltonian system constituted by the system of interest and two auxiliary ones controlling temperature and pressure, and the occurring of a process due to a change in the thermodynamic parameters is implemented by a suitable choice of the measure for the initial data.

I. INTRODUCTION

There are situations of metaequilibrium or metastability, in which the laws of phenomenological thermodynamics apply, whereas equilibrium statistical thermodynamics does not. Here we show how at least a first step can be made in providing a formulation of statistical thermodynamics for such cases, working in a classical Hamiltonian setting.

A paradigmatic case is that of partial (or incomplete) equilibrium, as Landau and Lifshitz call it (see 1, sec. 4). Macroscopically one has a system characterized by a phenomenological internal energy, say $U(T,p)$ if its thermodynamic state is defined by temperature T and pressure p. Microscopically one has a corresponding dynamical system with a set of degrees of freedom, a Hamiltonian H and a given family of invariant probability measures, for example the generalized Gibbs measure with parameters T and p. At equilibrium one makes the identification

$$U = \langle H \rangle_{T,p} \ ,$$

where $\langle \cdot \rangle_{T,p}$ denotes mean with respect to the invariant generalized Gibbs measure for the given values of T and p. In the case of partial (or incomplete) equilibrium, however, up to a given observation time τ only a fraction of the degrees of freedom is actually able to react to changes of the thermodynamic parameters.

Thus, considering for example the case of the specific heat at constant volume C_V, this is a well defined thermodynamic quantity which has a concretely measured value, but this value does not agree with the one predicted by equilibrium statistical mechanics. A typical case, mentioned by Landau and Lifshitz, is *"the partial equilibrium of a mixture of several substances which interact chemically. Owing to the relative slowness of chemical processes, the equilibrium connected with the motion of molecules occurs, in general, considerably sooner than the equilibrium connected with the interchange of atoms between the molecules, i.e., connected with the composition of the mixture"*. For example, in a system composed of a mole of graphite C and a mole of O_2, at room temperature combustion does not occur, and the measured heat capacity at constant volume is simply the sum of the specific heats of the two components, i.e., $1R + (5/2)R = (7/2)R$, whereas at equilibrium the heat capacity is that of one mole of carbon dioxide CO_2, namely, $(5/2)R$. Another situation presenting some analogy with partial equilibrium is that of the *"time dependent specific heat"* of Birge and Nagel [2].

On the other hand, in all the cases mentioned there exists a theoretical formula for C_V which explicitly depends on the observation time, namely, the formula provided by Fluctuation Dissipation (FD) theory, which is thus a good candidate for a theoretical counterpart to the phenomenological value. More in general, in situations of metaequilibrium one has available the differential forms δQ and δW of the heat absorbed by the system and of the work performed by it, with coefficients which have well defined phenomenological values. Analogously, the corresponding theoretical coefficients provided by FD theory have well defined values too, if a suitable dynamical re-

---

*Electronic address: andrea.carati@unimi.it
†Electronic address: alberto.mario.maiocchi@gmail.com
‡Electronic address: luigi.galgani@unimi.it

laxation did occur, albeit possibly a partial one.

However, there remains open the problem of proving that a consistent statistical thermodynamics can be formulated. At a phenomenological level, the internal energy can be defined by integration of the available differential form $\delta Q - \delta W$, which is closed in virtue of the first principle. At a theoretical level, instead, such a property should be proved, but a proof is apparently lacking.

In this paper we show that, in situations of metaequilibrium, the formulæ for the coefficients of the fundamental form $\delta Q - \delta W$ provided by FD theory can be expressed in such a way that the closure property of the form becomes manifest. This proves the first principle. Moreover, the second principle too is proven, as a consequence of microscopic time–reversibility. So one would have a complete statistical thermodynamics if one could also prove the validity of the zeroth principle (transitivity of metaequilibrium), or of a suitable generalization of it. In any case one has available a candidate for the internal energy, namely, the state function U which is obtained by integrating the form $\delta Q - \delta W$, with coefficients provided by FD theory. However, and this is perhaps the more relevant contribution of this paper, it turns out that in general such an internal energy U cannot be expressed as the mean of the Hamiltonian with respect to the invariant measure considered.

The main point thus consists in introducing a suitable definition for the statistical mechanical analogues of Q and W, as quantities parametrically dependent on the observation time τ (which we think of as fixed once for all, and will be mostly left understood). To this end we essentially follow the procedure used in FD theory, which in turn mimics the procedure of phenomenological thermodynamics, inasmuch as it takes into account only the fraction of energy which is actually exchanged with the measurement apparatus, when an increment is given to the thermodynamic parameters T and p of the latter. It is left to the dynamics itself, which involves the interactions between the system of interest and the measurement apparatus, to determine the amount of energy which is actually exchanged. It will be seen, however, that the definitions given here are not exactly those of FD theory. Some criticism was already raised against the latter [3]. The present definitions are in our opinion somehow simpler, and more suited to the aims of thermodynamics.

In order to implement the idea of mimicking the procedure used in phenomenological thermodynamics, we have first of all to consider the system of interest as coupled with two auxiliary ones, a thermostat which controls the temperature and a barostat which controls the pressure. The global system is studied as an isolated, time independent, Hamiltonian system. The existence of different states corresponding to different values of the macroscopic thermodynamic parameters T and p is taken into account by introducing different choices for the probability measure of the initial data in the global phase space. This way of handling the problem will make the proofs of the first and of the second principles particularly simple.

In section II the analogue of the fundamental form is defined and the first principle is proven. The second principle is proven in section III. To this end, making use of microscopic time–reversibility, we preliminarily establish formulæ of FD theory type, which express the coefficients of the form δQ in terms of equilibrium correlations of suitable functions. From this, in particular, the property $C_p \geq 0$ follows, which amounts to the classical formulation of the second principle in the form of Clausius. In section IV it is shown how the thermodynamic temperature of the system of interest coincides with that of the thermostat if a suitable dynamical relaxation takes place. Further discussions are given in the conclusive section.

In an appendix it is discussed how the present results may be of interest for the Fermi–Pasta–Ulam problem, the analogy of which with glassy–like systems was first proposed by a group of people about Parisi in the year 1982 [4] (see also 5 and the review 6).

II. THE MODEL, MICROSCOPIC ANALOGUE OF THE FUNDAMENTAL FORM, AND DEDUCTION OF THE FIRST PRINCIPLE

As usual (see 7, or 8, page 131), we take the system of interest (say, a fluid) enclosed in a cylindrical vessel one basis of which is fixed, while the other one is a movable piston. The distance of the piston from the fixed wall will be denoted by l, so that to any distance l there corresponds a volume $V = Al$ of the system of interest, A being the area of the piston. Through the two bases of the cylinder the system of interest can exchange energy with two other systems, a thermostat which acts as a source of heat (through the fixed basis), and a barostat which acts as a source of purely mechanical, or adiabatic, work (through the movable piston). It is well known [9] that in dynamical terms the adiabaticity condition is very well satisfied up to extremely long times if the piston is modeled as a single rigid particle of macroscopic mass. The two auxiliary systems are thought of as having no direct mutual interaction.

The global system (system of interest plus sources of heat and of work) is dealt with in a classical Hamiltonian setting, with a given time independent Hamiltonian and a fixed global geometry, the position of the piston being considered as one of the configurational coordinates of the global system. One thus has a global phase space with canonical coordinates

$$z = (x, x', x'', l, p_l)$$

where x are the canonical coordinates of the system of interest, x' those of the heat source, x'' those of the work source, l the distance of the piston from the fixed wall, and p_l the corresponding conjugate momentum. The total Hamiltonian is decomposed as

$$H_{\text{tot}}(z) = H(x, l) + H_{\text{ther}}(x') + H_{\text{bar}}(x'', l) + H_{\text{inter}}(z) .$$

Here H is the partial Hamiltonian of the system of interest, while H_{ther}, and H_{bar} are the partial Hamiltonians of the sources, the labels *ther* and *bar* standing for thermostat and barostat respectively. Finally there appears an interaction Hamiltonian, which should contain a term of the form $p_l^2/2M$ for the kinetic energy of the piston, dealt with as a single rigid particle of mass M of macroscopic size. As usual it is understood that, for large systems, the interaction Hamiltonian, although playing an essential dynamical role in allowing for energy exchanges, can be neglected in estimating mean values of the quantities of interest. Indeed such quantities are proportional to the volume of the system of interest, whereas the interaction terms are proportional to the area of the basis of the cylinder; furthermore, motions are considered in which the piston moves very slowly. The flow in the global phase space, induced by the global Hamiltonian H_{tot}, will be denoted by Φ^t.

Our aim is to estimate the energy exchanges between the system of interest and each of the sources (namely, the heat Q absorbed by the system and the work W performed by it). This involves first of all the dynamics of the global system, which induces corresponding exchanges for any initial datum z, and then the specification of the measure over the initial data which should be assigned in order to estimate mean values.

For what concerns the dynamics, along any solution of the equations of motion the total energy is conserved, so that, neglecting the contribution of the interaction Hamiltonian, one has energy conservation in the form

$$[H + H_{\text{ther}} + H_{\text{bar}}](\Phi^\tau z) = [H + H_{\text{ther}} + H_{\text{bar}}](z) \ ,$$

or equivalently

$$[\Delta H + \Delta H_{\text{ther}} + \Delta H_{\text{bar}}](z) = 0 \ , \qquad (1)$$

where we have introduced the notation

$$\Delta F(z) \equiv \Delta^\tau F(z) \stackrel{\text{def}}{=} F(\Phi^\tau z) - F(z) \qquad (2)$$

for the change of any dynamical variable F induced, at the observation time τ, by the time evolution Φ^t of the global system, as a function of the initial point z of phase space. Relation (1) will finally lead to a deduction of the first principle for the energy exchanges. This, however, requires to previously define the measure on the global phase space that should be used in estimating mean values, or even typicality of an orbit.

Clearly, invariant measures correspond to equilibrium states. Indeed the mean values of dynamical variables not explicitly depending on time, computed with respect to an invariant measure, turn out to be time independent. Thus the mean values of the three energies H, H_{ther} and H_{bar} do not change with time, i.e., there are no energy exchanges, notwithstanding the fact that no one of the three partial Hamiltonians is an integral of motion. Choosing a measure which is not invariant, leads instead to a thermodynamic process. We are interested

in processes induced by an increment of the thermodynamic parameters of the sources, with respect to those characterizing an invariant measure.

To this end we consider a family of measures relative to the global system. The family depends on two "external" thermodynamic parameters relative to the sources (say temperature T for the heat source and pressure p for the work source), and furthermore on at least two "internal" parameters, say $\alpha = (\alpha_1, \alpha_2)$, relative to the system of interest. So the family will have densities of the form

$$\rho(T, p, \alpha; z) \ .$$

We then assume that, for any pair T, p of external thermodynamic parameters, there exists an equilibrium state of the global system (that might be called *reference state*), which corresponds to some definite values $\overline{\alpha} = \overline{\alpha}(T, p)$ of the internal parameters α for which the measure is invariant. In particular the invariant measure could be a global Gibbs one. This is the typical case one considers in discussing partial equilibrium. The more general case considered above is of interest when discussing metastability.

Given one such equilibrium state with density

$$\rho(T, p, \overline{\alpha}; z) \ , \qquad (3)$$

we then introduce the state in which the external parameters have been incremented, *while the values of the internal parameters are kept constant*. In general such a measure will not be invariant, i.e., will correspond to a nonequilibrium situation, leading to a process. The corresponding measure will have a density

$$\rho(T + \mathrm{d}T, p + \mathrm{d}p, \overline{\alpha}; z) \ . \qquad (4)$$

Notice that, in the present model, a change of the external parameters can be obtained without any alteration of the global geometry of the system, because a change in the pressure p of the barostat can be produced by a suitable change of its temperature, which in turn can be described through a change of the measure on the initial data.

Mean values with respect to the noninvariant measure with density (4) will be denoted by $\langle \cdot \rangle^{\text{neq}}$, while those referring to the invariant measure with density (3) might be denoted by $\langle \cdot \rangle^{\text{eq}}$, and instead will be simply denoted by $\langle \cdot \rangle$:

$$\langle F \rangle^{\text{neq}} \stackrel{\text{def}}{=} \int F(z) \, \rho(T + \mathrm{d}T, p + \mathrm{d}p, \overline{\alpha}; z) \, \mathrm{d}z \ ,$$
$$\langle F \rangle \stackrel{\text{def}}{=} \int F(z) \, \rho(T, p, \overline{\alpha}; z) \, \mathrm{d}z \ . \qquad (5)$$

Let us now take as a measure for the initial data of the global system the noninvariant one with density (4). For any initial datum z, which evolves to $\Phi^\tau z$, there remain defined an increment of the energy of the thermostat and an increment of the energy of the barostat,

and this naturally leads to microscopic definitions of the heat Q absorbed by the system of interest and of the work W performed by it. Indeed, by analogy with the standard equilibrium procedure, for any given dynamical variable $F(z)$ one considers the increment corresponding to each initial datum, and then takes the mean value over the data distributed according to the given (nonequilibrium) measure. As usual, one might presume that such mean values are good representatives of the values actually taken in correspondence to single typical orbits.

So, recalling the meaning (2) of ΔF, we introduce the *definitions* which constitute the key point of the whole paper, namely,

$$Q \stackrel{\text{def}}{=} -\left\langle \Delta H_{\text{ther}} \right\rangle^{\text{neq}} \equiv$$
$$-\int \left[H_{\text{ther}}(\Phi^\tau z) - H_{\text{ther}}(z) \right] \rho(T + \mathrm{d}T, p + \mathrm{d}p, \overline{\alpha}; z)\, \mathrm{d}z \; ,$$
$$W \stackrel{\text{def}}{=} \left\langle \Delta H_{\text{bar}} \right\rangle^{\text{neq}} \; . \tag{6}$$

We can now come to the first principle. From the definitions (6) of Q and W, and from energy conservation (1), one gets

$$Q - W = \left\langle \Delta H(z) \right\rangle^{\text{neq}}$$
$$\equiv \int \Delta H(z)\, \rho(T + \mathrm{d}T, p + \mathrm{d}p, \overline{\alpha}; z)\, \mathrm{d}z \; . \tag{7}$$

Then, by expanding ρ to first order in $\mathrm{d}T$ and $\mathrm{d}p$, differential forms δQ and δW remain thus defined. So the analogue of the fundamental form $\delta Q - \delta W$ turns out to be expressed in terms of the Hamiltonian H of the system of interest only (although still in terms of a measure involving the two sources), being given by

$$\delta Q - \delta W = c_1(T, p)\, \mathrm{d}T + c_2(T, p)\, \mathrm{d}p \; , \tag{8}$$

with coefficients

$$c_1 \stackrel{\text{def}}{=} \int \Delta H(z)\, \frac{\partial}{\partial T}\rho(T, p, \overline{\alpha}; z)\mathrm{d}z \; ,$$
$$c_2 \stackrel{\text{def}}{=} \int \Delta H(z)\, \frac{\partial}{\partial p}\rho(T, p, \overline{\alpha}; z)\mathrm{d}z \; . \tag{9}$$

Proving the first principle amounts to proving that the differential form $\delta Q - \delta W$ is closed. On the other hand, the variables T and p enter the integrals at the right hand side of (9) only through the factor $\frac{\partial}{\partial T}\rho$ or the factor $\frac{\partial}{\partial p}\rho$. Thus, assuming that the density ρ is smooth enough in the parameters, one has

$$\frac{\partial c_1}{\partial p} = \frac{\partial c_2}{\partial T} \; ,$$

which is the analogue of a Maxwell relation. So the form $\delta Q - \delta W$ is closed, and there exists (at least locally) a function $U = U(T, p)$ such that $\delta Q - \delta W = \mathrm{d}U$. So the internal energy can be defined by integrating the differential form (8). The first principle is thus proven. Notice

that the proof is very simple due to the fact that we are considering a closed model, in which the changes in the external parameters involve only changes in the measures, and not in the dynamics (at variance with what occurs in standard FD theory).

We close this section with two remarks. The first one concerns the fact that the internal energy U cannot be expressed as the mean of the energy H with respect to the invariant measure. For example, the state function $\tilde{U}$ defined by

$$\tilde{U} = \int H(z)\, \rho(T, p, \overline{\alpha}; z)\mathrm{d}z$$

produces a differential form

$$\mathrm{d}\tilde{U} = \tilde{c}_1(T, p)\, \mathrm{d}T + \tilde{c}_2(T, p)\, \mathrm{d}p \; , \tag{10}$$

with coefficients

$$\tilde{c}_1 \stackrel{\text{def}}{=} \int H(z)\, \frac{\partial}{\partial T}\rho(T, p, \overline{\alpha}; z)\mathrm{d}z \; ,$$
$$\tilde{c}_2 \stackrel{\text{def}}{=} \int H(z)\, \frac{\partial}{\partial p}\rho(T, p, \overline{\alpha}; z)\mathrm{d}z \tag{11}$$

which, at variance with (9), involve integrals of H rather than of ΔH.

The second remark concerns the relation between the dynamical increment of internal energy

$$\Delta^\tau U \stackrel{\text{def}}{=} Q - W \tag{12}$$

defined through (7) – in correspondence to increments $\mathrm{d}T$, $\mathrm{d}p$ of the parameters – and the increment of the equilibrium internal energy which is obtained by incrementing the parameters T and p in the equilibrium measure. This relation is obtained by remarking that the definition (7), which involves the time evolved $H(\Phi^\tau z)$ of the dynamical variable H while keeping fixed the measure, can also be dually expressed by letting the measure evolve with time, while keeping fixed the dynamical variable H:

$$\Delta^\tau U = \int H(z) \Big[\rho(T + \mathrm{d}T, p + \mathrm{d}p, \overline{\alpha}; \Phi^{-\tau} z)$$
$$- \rho(T + \mathrm{d}T, p + \mathrm{d}p, \overline{\alpha}; z) \Big] \mathrm{d}z \; . \tag{13}$$

Thus, if the density ρ tends (in weak sense) to the final equilibrium density, with the internal parameters α adapted to the new external ones, and if in addition it factors into an external part and an internal one (as should be assumed of the initial density too), then one should have $\Delta^\tau U \to \left\langle H \right\rangle_{T+\mathrm{d}T, p+\mathrm{d}p} - \left\langle H \right\rangle_{T,p}$, i.e.,

$$\left\langle H \right\rangle_{T,p} + \Delta^\tau U \to \left\langle H \right\rangle_{T+\mathrm{d}T, p+\mathrm{d}p} \; ,$$

where $\left\langle \cdot \right\rangle_{T,p}$ denotes mean value over the phase space of the system of interest, with respect to the generalized Gibbs measure at values p, T of the parameters.

So the time–dependent family $\Delta^\tau U$ considered here provides an interpolation $\langle H \rangle_{T,p} + \Delta^\tau U$ between the initial equilibrium value $\langle H \rangle_{T,p}$ and the final one $\langle H \rangle_{T+\mathrm{d}T,p+\mathrm{d}p}$ of the internal energy. However, it is left to the dynamics to decide whether, and at which time, will the final equilibrium value be actually attained.

III. THE SECOND PRINCIPLE AS A CONSEQUENCE OF MICROSCOPIC TIME–REVERSIBILITY.

The result proved in the previous section, according to which in general the internal energy of metaequilibrium thermodynamics cannot be expressed as the mean of the Hamiltonian with respect to the invariant measure, is the one to which we attach a particular significance, as will be discussed in the appendix. In this section and in the next one, however, we continue the discussion of metaequilibrium thermodynamics. Here we show how the second principle too is obtained. To this end we will show that the coefficients of the fundamental form $\delta Q - \delta W$ can be expressed in terms of suitable correlation functions, as in FD theory. This fact is obtained if the further assumptions are made that i) the dynamics is time reversible, and that ii) the reservoirs are described by Gibbs measures. The latter condition can actually be weakened, as will be mentioned later.

So, let us recall the time–reversibility property. This property amounts to requiring that there exists a mapping $\mathcal{P}$ of the phase space onto itself (inverting the sign of the velocities of all particles that constitute the global system) with the properties

$$\mathcal{P}^2 = \mathcal{I} \ , \quad (\mathcal{P}\Phi^\tau)^2 = \mathcal{I} \ ,$$

where $\mathcal{I}$ is the identity map. Obviously the invariance will be required to hold also with respect to the considered measure. It is well known that the reversibility property is satisfied by essentially all systems of interest, namely, those with Hamiltonians even in the momenta of the particles. So in particular in our case we will have

$$H_{\mathrm{tot}}(\mathcal{P}z) = H_{tot}(z) \ , \quad H(\mathcal{P}z) = H(z) \ ,$$
$$H_{\mathrm{ther}}(\mathcal{P}z) = H_{\mathrm{ther}}(z) \ , \quad \dots \quad .$$

Concerning the family of measures for the initial data, we will assume that the invariant measures have the form

$$
\begin{aligned}
&\rho(\beta, \beta_{\mathrm{bar}}, \overline{\alpha}, z) \\
&= c(\beta, \beta_{\mathrm{bar}}) \ e^{-\beta H_{\mathrm{ther}}} \ e^{-\beta_{\mathrm{bar}} H_{\mathrm{bar}}} \ \rho^{\mathrm{syst}}(\overline{\alpha}; z) \ ,
\end{aligned}
\tag{14}
$$

where ρ^{syst} refers to the system of interest, $c(\beta, \beta_{\mathrm{bar}})$ is the familiar normalization factor involving the partition functions of the reservoirs, while β and β_{bar} are Lagrange multipliers which determine the temperature T of the thermostat and the temperature of the barostat (and so implicitly its pressure too).

Our aim is now to find a suitable expression for W and thus for Q. First of all the equilibrium pressure p is defined as usual (see 8) as the equilibrium mean value of the corresponding dynamical variable $P(z) \stackrel{\mathrm{def}}{=} \frac{\partial H}{\partial V}(z) = \frac{1}{A}\frac{\partial H}{\partial l}(z)$. So the pressure p is given by

$$p = \langle P \rangle = \left\langle \frac{\partial H}{\partial V} \right\rangle \ .$$

Thus the work performed by the system of interest on the barostat should be given by

$$W = p\langle \Delta V \rangle^{\mathrm{neq}} \ . \tag{15}$$

So the heat Q absorbed by the system of interest, $Q = \langle \Delta H \rangle^{\mathrm{neq}} + W$ (see (7)), will eventually be expressed in terms of dynamical variables of the system of interest, as

$$Q = \langle \Delta(H + pV) \rangle^{\mathrm{neq}} \ .$$

We thus turn to investigating the expression for the specific heat. We even perform the more general computation which gives the differential form δQ in terms of $\mathrm{d}\beta$ and $\mathrm{d}\beta_{\mathrm{bar}}$. Such a form is obtained from Q through the same procedure used in the previous section for the fundamental form $\delta Q - \delta W$, i.e., by expanding the nonequilibrium density to first order in $\mathrm{d}\beta$ and $\mathrm{d}\beta_{\mathrm{bar}}$. So one gets

$$\delta Q = q_1(\beta, \beta_{\mathrm{bar}}) \, \mathrm{d}\beta + q_2(\beta, \beta_{\mathrm{bar}}) \, \mathrm{d}\beta_{\mathrm{bar}} \ , \tag{16}$$

where the coefficients are given by

$$
\begin{aligned}
q_1 &\stackrel{\mathrm{def}}{=} \int \Delta(H + pV)(z) \ \frac{\partial}{\partial \beta} \rho(\beta, \beta_{\mathrm{bar}}, \overline{\alpha}; z)\mathrm{d}z \ , \\
q_2 &\stackrel{\mathrm{def}}{=} \int \Delta(H + pV)(z) \ \frac{\partial}{\partial \beta_{\mathrm{bar}}} \rho(\beta, \beta_{\mathrm{bar}}, \overline{\alpha}; z)\mathrm{d}z \ .
\end{aligned}
\tag{17}
$$

One now has the following
Lemma: *Let the global Hamiltonian system be time–reversible, and let the invariant measure have a density of the form (14). Then, for any dynamical variable F having the property*

$$F(\mathcal{P}z) = F(z) \ , \tag{18}$$

one has

$$
\begin{aligned}
\int \Delta F \ \frac{\partial}{\partial \beta} \rho \mathrm{d}z &= -\frac{1}{2} \left\langle \Delta F \ \Delta(H + pV) \right\rangle \\
\int \Delta F \ \frac{\partial}{\partial \beta_{\mathrm{bar}}} \rho \mathrm{d}z &= \frac{p}{2} \left\langle \Delta F \ \Delta V \right\rangle \ .
\end{aligned}
\tag{19}
$$

The proof, for example for the first formula, is as follows. Due to the exponential form of ρ with respect to the reservoirs one has

$$I \stackrel{\mathrm{def}}{=} \frac{\partial}{\partial \beta} \langle \Delta F \rangle = -\int \Delta F(z) H_{\mathrm{ther}}(z)\rho(z)\mathrm{d}z \ .$$

Notice that, in performing the partial differentiation, one further term would appear, which contains the derivative

of the normalization factor $c(\beta, \beta_{\text{bar}})$. However, being proportional to $\langle \Delta F \rangle$, such a further term vanishes, in virtue of the invariance of the considered measure. Introduce then the change of variables

$$z = \mathcal{P}\Phi^\tau y$$

which is canonical, so that $\mathrm{d}z = \mathrm{d}y$. Use (18), which implies

$$\Delta F(z) = -\Delta F(y) ,$$

and use also the invariance of the measure, $\rho(z) = \rho(y)$. Thus, calling again by z the dummy variable y one gets

$$I = \int \Delta F(z) H_{\text{ther}}(\mathcal{P}\Phi^\tau z)\rho(z)\mathrm{d}z .$$

Finally, the result follows by taking for I the semisum of the original expression and of the last one, using

$$\Delta H_{\text{ther}} = -\Delta(H + pV) .$$

The second formula is proved analogously, using $\Delta H_{\text{bar}} = p\,\Delta V$.

Thus the coefficients q_1, q_2 of the form $\mathrm{d}Q$ turn out to be expressed in terms of correlation functions with respect to the invariant measure:

$$
\begin{aligned}
q_1 &= -\frac{1}{2}\left\langle \left[\Delta(H + pV)\right]^2 \right\rangle \\
q_2 &= \frac{p}{2}\left\langle \Delta(H + pV)\,\Delta V \right\rangle .
\end{aligned}
\tag{20}
$$

In particular, using $\frac{\partial}{\partial T} = -(1/(k_B T^2)\frac{\partial}{\partial\beta}$ one finds that the specific heat C_p at constant β_{bar} (i.e., at constant p), is manifestly positive,

$$C_p = \frac{1}{2k_B T^2}\left\langle \left[\Delta(H + pV)\right]^2 \right\rangle \geq 0 . \tag{21}$$

Thus the second principle in the Clausius form (heat spontaneously flows from hot to cold bodies) is proven. By the way, one is naturally led to guess that the second principle may not hold if microscopic time–reversibility were not satisfied.

The positiveness of the specific heat was proven above, making use of the assumption that the marginal measure for the environment be of Gibbs type. However, if one looks at the proof one realizes that the result also holds if any assumption is made which guarantees that the partial derivatives of the density ρ have negative definite sign. For example, a sufficient condition is that, in formula (14) for the density, instead of the exponential $\exp(-\beta H)$ there appears a factor $f(\beta H)$, with any function f having the property $f' \leq 0$.

Having thus proven positiveness of the specific heat, even in such a more general frame, the existence of an integrating denominator for the form δQ then follows from arguments of Carathéodory type. A connection with the temperature of the thermostat will be provided in the next section.

Finally, the fact that the thermodynamic quantities may attain the corresponding equilibrium values if a suitable relaxation takes place, is exhibited by expressing the thermodynamic coefficients in terms of time correlation functions. For example, in the case of the specific heat C_p given by (21), by expanding the square

$$[\Delta(H + pV)(z)]^2 = [(H + pV)(\Phi^\tau z) - (H + pV)(z)]^2$$

and adding and subtracting the term $\langle H + pV \rangle^2$ one immediately gets

$$C_p = \frac{1}{k_B T^2}\left[\sigma^2_{H+pV} - \mathscr{C}_{H+pV}(\tau)\right] , \tag{22}$$

where σ^2_F denotes the variance of F and $\mathscr{C}_F(\tau) \overset{\text{def}}{=} \langle F(\Phi^\tau z)F(z) \rangle - \langle F \rangle^2$ its time–autocorrelation function.

IV. THE TEMPERATURE OF THE THERMOSTAT AS AN INTEGRATING DENOMINATOR FOR δQ

Thus, having proven the second principle, we know that an integrating denominator for the differential form δQ exists. One may ask however whether such an integrating denominator coincides with the temperature $1/\beta$ of the thermostat. We will show now that this is actually the case, at least if the observation time τ is sufficiently large to guarantee that a suitable dynamical decorrelation did occur.

So, let us consider the form

$$\beta\delta Q = \beta q_1\,\mathrm{d}\beta + \beta q_2\,\mathrm{d}\beta_{\text{bar}} , \tag{23}$$

where the coefficients q_i are defined by (17), and let us investigate whether it is closed. This requires to calculate $\frac{\partial(\beta q_1)}{\partial\beta_{\text{bar}}} - \frac{\partial(\beta q_2)}{\partial\beta}$. As in the case of the internal energy, again one finds that the two terms involving the second derivatives of ρ cancel, so that one remains with

$$
\begin{aligned}
&\frac{\partial(\beta q_1)}{\partial\beta_{\text{bar}}} - \frac{\partial(\beta q_2)}{\partial\beta} \\
&= \beta\frac{\partial p}{\partial\beta_{\text{bar}}}\int \Delta V(z)\,\frac{\partial}{\partial\beta}\rho\,\mathrm{d}z \\
&\quad - \int \Delta(H + pV)(z)\,\frac{\partial}{\partial\beta_{\text{bar}}}\rho\,\mathrm{d}z .
\end{aligned}
\tag{24}
$$

So, using the Lemma one gets

$$
\begin{aligned}
&\left[\frac{\partial(\beta q_1)}{\partial\beta_{\text{bar}}} - \frac{\partial(\beta q_2)}{\partial\beta}\right] \\
&= -\frac{1}{2}\left\langle \Delta V\,\Delta(H + pV) \right\rangle\left(\beta\frac{\partial p}{\partial\beta_{\text{bar}}} + p\right) .
\end{aligned}
\tag{25}
$$

Thus the form is closed if one has

$$\frac{\partial p}{\partial\beta_{\text{bar}}} = -\frac{p}{\beta} . \tag{26}$$

To investigate this point requires to find the expression of $\frac{\partial p}{\partial \beta_{\mathrm{bar}}}$, and this is obtained by finding the expression of dp as a differential form with respect to $d\beta$, $d\beta_{\mathrm{bar}}$. In analogy with what previously done for other differential forms, we just define dp as the first order expansion of $\langle \Delta P \rangle^{\mathrm{neq}}$ with respect to $d\beta$, $d\beta_{\mathrm{bar}}$. This gives, using again the Lemma,

$$dp = -\frac{1}{2}\langle \Delta P \Delta (H + pV)\rangle\, d\beta + \frac{p}{2}\langle \Delta P \Delta V\rangle\, d\beta_{\mathrm{bar}} \quad (27)$$

from which one gets

$$\frac{\partial p}{\partial \beta_{\mathrm{bar}}} = \frac{p}{2}\langle \Delta P \Delta V\rangle\,. \quad (28)$$

On the other hand, if a suitable decorrelation occurs, one has as usual

$$\frac{1}{2}\langle \Delta P \Delta V\rangle \to \langle \left(P - \langle P\rangle\right)\left(V - \langle V\rangle\right)\rangle\,. \quad (29)$$

Finally, the r.h.s. of (29) is just equal to $-1/\beta$, as one finds in the Landau–Lifshitz textbook [1].

Thus after a sufficiently long time the temperature of the thermostat is an integrating denominator of the form δQ. It is presumable however that such a relaxation time might in general be rather short, especially in comparison with the relaxation times related to the specific heats at low temperatures.

V. CONCLUSIONS

In this paper we have shown how it is possible to formulate statistical thermodynamics for a Hamiltonian system, in terms of the processes induced by a change of the thermodynamic parameters. A characteristic feature of our approach is that it allows one to deal with metaequilibrium (or partial equilibrium) situations in which the time–correlations did not decay to zero, in addition to the equilibrium situations described in the familiar way through equilibrium ensembles. An approach somehow similar to ours was recently proposed by a C. Bernardin and S. Olla [10]. The main difference is that we frame the problem in a time–independent Hamiltonian setting in which the processes are induced by nonequilibrium measures for the initial data, whereas in the quoted paper the processes are induced by imposing suitable time dependencies on the parameters of the global system.

A relevant feature of the present approach is that it involves in a substantial way, in addition to the system of interest, also the measurement apparatuses. Consequently, one has the problem of understanding whether a physical, actually measured, property belongs to the system of interest or rather depends on the measuring procedure too. Standard equilibrium statistical thermodynamics, being formulated in terms of the Hamiltonian of the system of interest only, avoids in principle any reference to the method of measurement, and perhaps one would be tempted to say that this is indeed correct. In general, however, things are not so. And actually, in the work 2, of which it was said that it gave rise to a new epoch of thermodynamics, one even finds the sentence: *"How does one interpret a quantity such as the specific heat in a nonequilibrium situation? Clearly, the dynamics of how one performs the measurement enters into the measured value"*. This circumstance finds a counterpart in the present approach. Indeed the thermodynamic quantities we are considering, typically the specific heat of the system of interest, are in general represented by expressions which, through the intermediate of the measure, involve not only the system of interest, but also the sources. So the measured quantities in general depend on the measurement procedure, and the dependence disappears only when a suitable relaxation has occurred, if indeed it takes place.

However, the main interest of the proposal made here is that it may provide an answer to a question of principle that remains open within the equilibrium approach. The problem is to understand how many degrees of freedom should be attributed to a microscopic model. Why can one neglect the rotational degrees of freedom in a system of hard, completely smooth, spheres? Conversely, why does a supercooled liquid apparently have a number of degrees of freedom larger than the corresponding crystal?

Problems of this type are usually met in chemistry or in glasses (see for example 11). In fact, however, they touch on questions of principle that afflicted very much Boltzmann (see 12 and the quotations reported in 13) and Gibbs [14], and were later taken up by authors such as Rayleigh and Jeans (see the first Solvay Conference [15], particularly the letter sent by Rayleigh) and Nernst [16], in connection with foundational problems of quantum mechanics.

In an appendix it is discussed how the present results may be of interest for the Fermi–Pasta–Ulam problem, which also presents metaequilibrium situations (as apparently first suggested by Parisi and collaborators[4]), and thus constitutes a modern arena for dealing with foundational problems of the type just mentioned.

Appendix

Possible Relevance of Metaequilibrium Thermodynamics for the Fermi–Pasta–Ulam Problem. In this paper we have shown how a statistical thermodynamics different from the equilibrium one can be formulated in situations in which the system does not present a full dynamical relaxation, i.e., in situations of partial equilibrium, as Landau and Lifshitz would say. This can occur even if the system was originally in an equilibrium state, described by an invariant Gibbs measure. In particular, if the dynamics is such that the decorrelation of $H + pV$ is negligible for long times when $T \to 0$, i.e., if one has $\mathscr{C}_{H+pV}(\tau) \simeq \sigma^2_{H+pV}$ for long times when $T \to 0$, then one has $C_p \simeq 0$, and so the system behaves in agree-

ment with the third principle.

Situations of this type were conceived by Boltzmann since his celebrated paper published in the journal Nature[12] in the year 1895 (see also the quotations reported in 13) and were amply discussed in the year 1916 by Walther Nernst[16], who came to the point of conceiving that equipartition of energy might be compatible with Planck's law. According to Nernst, the key point consists in distinguishing between the mechanical energy (which corresponds to equipartition if the initial data are Gibbs distributed), and the energy that can be dynamically exchanged in a process. Only the exchangeable energy should be identified with the thermodynamic internal energy U, while the complementary part should play the role of a kind of "zero–point energy".

Foundational questions of this type were reconsidered later in the context of the FPU problem. As previously mentioned, the idea that in the FPU problem one meets with metaequilibrium situations analogous to those of glasses was first advanced by a group of people about Parisi in the year 1982 [4] (see also 5 and the review 6).

In the FPU problem one has a classical approach, in which initial data are considered with only some of the normal modes excited (typically, the low frequency ones), and the result found is that, for low enough specific energies, up to extremely long times energy remains confined among a certain group of low frequency modes. Already in the year 1972, in a paper of Cercignani et al. [17] by the title *"Zero–point energy in classical non–linear mechanics"* it was however proposed that even in the FPU problem one might meet with situations of the type conceived by Nernst.

Numerical studies on the variant of the FPU problem with Gibbs distributed initial data came later, with results that were variously interpreted (see 18, 19). Finally a case was considered, which involves a situation very similar to the one discussed here [20], where one measures the specific heat of an FPU system (with an initial Gibbs distribution), put in contact with a heat reservoir. The result found is that at high enough energies a complete decorrelation occurs, and the standard equilibrium value is measured. Instead, at low enough energies situations of an apparent metaequilibrium (or partial equilibrium) are met, and smaller values of the specific heat are measured. Situations of metastability with hysteresis effects had been previously reported [21]. Finally, the time autocorrelation function of the mode energies, or of the energies of packets of nearby frequencies, were numerically studied in the paper 22.

So much for what concerns numerical results. In the meantime a consistent progress was achieved at an analytical level, because it was proven that a weak version of perturbation theory can be formulated in the thermodynamic limit [23]. At variance with the classical approach, one renounces to control all initial data, and only looks at quantities such as the time–autocorrelation functions of suitable dynamical variables, estimated with respect to the Gibbs measure. This allows one to obtain results which hold in the thermodynamic limit of infinitely many particles, with a finite temperature. It was found that such autocorrelation functions do not decay to zero up to rather long times: a Nekhoroshev–like exponential stability was found for a Φ^4–type model [24], while one step of perturbation theory was performed in the very FPU model [25]. Such results appear to disprove the common expectation (see for example 10) that situations of partial equilibrium as those discussed here should not occur for infinite systems.

Concerning the connections with the foundations of quantum mechanics, we now add a comment on the work of Poincaré of the year 1912 on the necessity of quantization [26] (see also the work 27 of Ehrenfest). The problem was posed in the following way. Assume that the internal energy U of the system of interest (the blackbody) be expressed as the mean value of the Hamiltonian H with respect to a suitable measure,

$$U = \int H(z)\,\sigma(z)\,\mathrm{d}z \ . \tag{30}$$

Then, it is asked which qualitative properties should the measure σ have in order that the function U given by (30) agree with the phenomenological internal energy. The answer of Poincaré was that, under quite general assumptions, quantization is necessary, i.e., the measure has to be singular.

Our comment is that, if one takes the metaequilibrium perspective, the claim that U should be expressed as an equilibrium mean value of the energy H (rather than a nonequilibrium mean value of $\Delta^\tau H$), is unfounded. The result of Poincaré shows that, if one does so, i.e., if one insists in representing U as an equilibrium mean value of H, then everything goes as if energy were quantized. Or, in Einstein's words at his contribution to the first Solvay Conference (see 28, section 3), *"The statistical properties of the phenomenon are the same as if energy were transferred only by integer quanta of size hν"*.

Curiously enough, an opinion similar to the latter one of Einstein was expressed by Poincaré too, in a paper of a less technical character [29] written shortly after the previously mentioned one, where he says: *"Will discontinuity reign over the physical universe and will its triumph be definitive? Or rather will it be recognized that such a discontinuity is only an appearance and that it dissimulates a series of continuous processes? The first person that saw a collision believed to be observing a discontinuous phenomenon, although we know today that the person was actually seeing the effect of very rapid changes of velocity, yet continuous ones."* And the conclusion was: *"To try to express today an opinion about these problems would mean to be wasting one's ink."*

Now, our admiration for Poincaré is unlimited, but our personal feeling, or rather hope, is that perhaps on this point Einstein was seeing farther than him. It appears that the main intuition of Einstein about the role of fluctuations was correct, although *"rather ironically"* (to use the words of 3) he couldn't have a full appreciation of

their dynamical features, which was obtained with FD theory only half a century later. It is just the stabilization of the time–autocorrelation functions to well definite nonvanishing values that allows for a thermodynamics different from the equilibrium one to hold. And this, not only for chemical systems or for glassy–like ones, but perhaps also for systems of FPU type, which are of interest

for the foundations of physics. By the way, if one looks at the objections raised by Poincaré against the "partial or incomplete equilibrium" point of view of Rayleigh and Jeans, one realizes that what he could not conceive is just that such partial equilibrium situations may actually occur in nature, and that they may be dealt with through the methods of phenolemological thermodynamics.

[1] L.D. Landau, E.M. Lifshitz, *Statistical Physics*, Pergamon Press (London, 1958).
[2] N.O. Birge, Phys. Rev. B **34**, 1631 (1986); N.O. Birge, S.R. Nagel, Phys. Rev. Lett. **54**, 2674 (1985).
[3] J.K. Nielsen, J.C. Dyre, Phys. Rev. B **54**, 15754 (1996).
[4] E. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo, A. Vulpiani, J. Phys. **43**, 707 (1982).
[5] G. Parisi, Europhys. Lett. **40**, 357 (1997).
[6] B. Benettin, A. Carati, L. Galgani, A. Giorgilli, in *The FPU problem: a status report*, G. Gallavotti ed., Lecture Notes in Physics Vol. 728, Springer Verlag (Berlin, 2007).
[7] E. Fermi, *Thermodynamics*, Dover (New York, 1956).
[8] A.J. Khinchin, Mathematical Foundations of Statistical Mechanics, Dover (New York, 1949).
[9] N.I. Chernov, J.L. Lebowitz, Ya.G. Sinai, Uspekhi Mat. Nauk **57** (6), 3 (2002).
[10] C. Bernardin, S. Olla, *Non–equilibrium macroscopic dynamics of chains of anharmonic oscillators*, preprint.
[11] M. Mézard, G. Parisi, Phys. Rev. Lett. **82**, 747 (1999).
[12] L. Boltzmann, Nature **51**, 413 (1895).
[13] A. Carati, L. Galgani, B. Pozzi, *The problem of the rates of thermalization and the relations between classical and quantum mechanics.* in M. Fabrizio et al. eds., *Mathematical models and methods for smart materials*, Advances in Mathematics n. 62, World Scientific (Singapore, 2002). Also available at Luigi Galgani's homepage.
[14] J.W. Gibbs, Elementary Principles in Statistical Mechanics, Dover (New York, 1960).
[15] *La théorie du rayonnement et les quanta*, P. Langevin and M. de Broglie eds., Gauthier–Villard (Paris, 1912).
[16] W. Nernst, Verhandl. Deutschen Phys. Gesell. **18**, 83 (1916).
[17] C. Cercignani, L. Galgani, A. Scotti, Phys. Lett. **38A**, 403 (1972); L. Galgani, A. Scotti, Phys. Rev. Lett. **28**, 1173 (1972).
[18] R. Livi, M. Pettini, S. Ruffo, A. Vulpiani, J. Stat. Phys. **48**, 539 (1987).
[19] A. Perronace, A. Tenenbaum, Phys. Rev. E **57**, 100 (1998); G. Marcelli, A. Tenenbaum, Phys. Rev E **68**, 041112 (2003).
[20] A. Carati, L. Galgani, Europhys. Lett. **74**, 528 (2006).
[21] A. Carati, L. Galgani, J. Stat. Phys. **94**, 859 (1999).
[22] A. Carati, L. Galgani, A. Giorgilli, S. Paleari, Phys. Rev. E **76**, 022104 (2007).
[23] A. Carati. J. Stat. Phys. **128**, 1057 (2007).
[24] A. Carati, A. Maiocchi, Comm. Math. Phys. **314**, 129 (2012).
[25] D. Bambusi, A. Carati, A. Maiocchi, in preparation.
[26] H. Poincaré, J. Phys. Th. Appl. **2**, 5 (1912); Œuvres IX, 626.
[27] P. Ehrenfest Ann. der Phys. **36**, 103 (1911).
[28] A. Einstein, *L'état actuel du problème des chaleurs spécifiques*, Contribution to the first Solvay Conference.
[29] H. Poincaré, Revue Scientifique **17**, 225 (1912); Œuvres IX, 654.